processing enzyme system. Partial processing also occurred at one or both dibasic sequences between B and A chains within proinsulin and other insulin precursors containing a short spacer peptide (containing 6 or more amino acid residues) in place of the C peptide. In contrast, no processing was observed in the absence of a spacer peptide in the insulin precursor molecule, e.g. B-Arg-Arg-A (where A and B are the A and B chain of human proinsulin, respectively). This type of single-chain insulin precursors could enzymatically be converted into insulin by treatment with trypsin and carboxypeptidase B.

Diers et al., *Drug Biotechnology Regulations (Scientific Basis and Practices),* Chiu and Gueriguian, Eds., Marcel Dekker, Inc., New York, 1991, pp. 167-177, describe the unfolded peptide as a leader or prosegment, next a Lys-Arg sequence, the B chain (amino acids 1-29), a short peptide bridge, followed by the A chain (amino acids 1-21). In this precursor, amino acid 29 of the B chain of insulin is connected to amino acid 1 of the A chain by a short connecting peptide containing one basic amino acid adjacent to the A chain.

Human insulin is produced through transpeptidation followed by hydrolysis of the ester bond formed. Several chromatography steps follow for further purification.

1.2.5. FOLDING OF INSULIN PRECURSORS

Human insulin is a protein possessing two amino acid chains of 51 amino acid residues in all. Six cysteine residues are present in the two amino acid chains, which in each case two cysteine residues being linked to each via a disulfide bond. Statistically, there are 15 possibilities of forming disulfide bridges within one human insulin molecule. However, only one of the 15 possibilities exists in biologically active human insulin with the following disulfide bridges: 1) A6-A11; 2) A7-B7; and 3) A20-B19.

The formation of the disulfide bridges which are present in human insulin is effected by way of an intermediate, with the cysteine residues of the human insulin being provided with a sulfur protective group, e.g., with a S-sulfonate (-S-SO_3^-) group (EP 0,037,255). In addition, pig proinsulin in which the cysteine residues are present as thio residues (-SH) has been used to obtain proinsulin possessing correctly linked cysteine bridges (*Biochemistry*, 1968, 60:622-629). Obermeier et al. described a process for obtaining proinsulin possessing correctly linked cysteine bridges from a corresponding proinsulin amino acid chain at a concentration of 0.05 to 0.3 g per liter in the presence of mercaptan, chaotropic auxiliary agents and hydrophobic absorber resins (U.S. Patent No. 5,473,049). The step of oxidative sulfitolysis is eliminated in the process described in U.S. Patent No. 5,473,049. However, the insulin protein can only be folded at a low concentration, which greatly diminishes the commercial value of this process. In addition, the use of large amount of mercaptan and hydrophobic absorber resins increase process

complexity and down-stream purification costs. From the disclosure of the U.S. Patent No. 5,473,049, it is unclear whether the benefit of eliminating the step of oxidative sulfitolysis step will outweigh the increased down-stream purification costs.

Citation of references hereinabove shall not be construed as an admission that such references are prior art to the present invention.

2. SUMMARY OF THE INVENTION

The present invention relates to a chimeric protein comprising, from N-terminus to C-terminus: a) a first peptidyl fragment consisting of an amino acid sequence that has at least 40% identity to a domain containing at least first 20 N-terminal amino acids of human growth hormone (hGH) protein, in which the percentage identity is determined over an amino acid sequence of identical size to the domain of hGH; b) an Arg residue, or a Lys residue, or a second peptidyl fragment consisting of at least 2 amino acids in which peptidyl fragment the most C-terminal amino acid residue is an Arg or a Lys; and c) a third peptidyl fragment consisting of an amino acid sequence containing more than two cysteine (Cys) residues which peptidyl fragment is not a portion of hGH protein. In particular, the invention relates to a chimeric protein wherein the third peptidyl fragment is an insulin precursor.

The invention also relates to a process for obtaining a first correctly folded insulin-precursor-containing chimeric protein comprising contacting an incorrectly folded second insulin-precursor-containing chimeric protein, which said second insulin-precursor-containing chimeric protein consists of an intramolecular chaperone (IMC) like peptidyl fragment separated from the insulin precursor by one or more cleavable amino acid residues, with at least one chaotropic auxiliary agent in an aqueous medium; wherein said IMC like peptidyl fragment: a) contains from about 20 to about 200 amino acid residues; b) is not the insulin precursor or a portion thereof; and c) improves the insulin precursor folding such that the yield of the correctly folded first insulin-precursor-containing chimeric protein when the incorrectly folded second insulin-precursor-containing chimeric protein is contacted with the chaotropic auxiliary agent is higher than the yield of the correctly folded insulin precursor when the incorrectly folded insulin precursor, which does not contain said IMC like peptidyl fragment, is contacted with the same chaotropic auxiliary agent.

The present invention further relates to an assay for screening an amino acid sequence for the ability to improve folding of an insulin precursor, comprising: (a) changing the amino acid sequence of the first peptidyl fragment of a chimeric protein disclosed in Section 4.2, which contains an insulin precursor, obtaining said chimeric

protein with said changes, contacting said chimeric protein with said changes with at least one chaotropic auxiliary agent in an aqueous medium under conditions and for a time sufficient such that said chimeric protein folds correctly, and measuring the folding yield of said chimeric protein with said changes; (b) obtaining the chimeric protein used in step (a), but without any amino acid sequence changes described in step (a), contacting the chimeric protein without any amino acid sequence changes described in step (a) with the same chaotropic auxiliary agents used in step (a) in an aqueous medium under the same conditions and for a same time used in step (a), and measuring the folding yield of the chimeric protein; and (c) comparing the folding yield of the chimeric proteins measured in step (a) and (b), respectively, in which the yield measured in step (a) substantially equals or is greater than the yield measured in step (b) indicates that the amino acid sequence improves folding of the insulin precursor.

3. BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B. Structure of proinsulin and mature insulin with correctly formed disulfide bridges. 1A depicts the structure of proinsulin. 1B depicts the structure of mature insulin with correctly formed disulfide bridges.

Figure 2. Map of the hGH-mini-proinsulin (SEQ ID NO:6) expression vector (pZRhi-1).

4. DETAILED DESCRIPTION OF THE INVENTION

Recombinant processes make it possible to produce human proinsulin, or proinsulin with an amino acid sequence and/or amino acid chain length diverging from that of a natural human insulin, in microorganisms. One major problem in the production of human proinsulin or its derivatives in microorganisms such as *E. coli* is the intracellular degradation (Ladisch and Kohlmann, *Biotechnol. Prog.*, 1992, 2:469478). In addition, human proinsulin or its derivatives recombinantly produced in microorganisms do not possess correctly linked cysteine bridges (U.S. Patent No. 5,473,049).

Prior to the present invention, a known process for obtaining human insulin recombinantly is based on the following procedures: 1) fermentation of the microorganisms transformed with a vector expressing a fusion protein containing human proinsulin or its derivatives; 2) cell disruption; 3) isolation of the fusion protein; 4) cleavage of the fusion protein with cyanogen bromide; 5) isolation of the cleavage product having the proinsulin sequence; 6) oxidative sulfitolysis; 7) formation of the correctly linked cysteine bridges; 8) desalting of the proinsulin; 9) chromatographic purification of the proinsulin possessing the